DOCUMENT RESUME

ED 355 905 IR 015 964

AUTHOR	Perkins, Bob
TITLE	Differences between Computer Administered and Paper Administered Computer Anxiety and Performance Measures.
PUB DATE	Feb 93
NOTE	14p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Anxiety; Comparative Analysis; *Computer Assisted Instruction; *Computer Assisted Testing; Computer Science Education; Graduate Students; Higher Education; Hypermedia; Inservice Teacher Education; *Performance Factors; Predictor Variables; Preservice Teacher Education; Pretests Posttests; Sex Differences; Student Attitudes; *Test Format; Undergraduate Students
IDENTIFIERS	*Computer Anxiety; *Paper and Pencil Tests

ABSTRACT

This research investigated whether computer anxiety is different if the measure is administered by computer rather than by paper and pencil. The study compared two groups of students (N=83) who were gathered from three undergraduate sections and one graduate section of a required computer class for in-service and pre-service teachers using anxiety level and performance as measures. Both groups took a written pretest; but one group was administered an anxiety scale (pre and post) and posttest on the computer using a HyperCard stack, while the other group used a paper and pencil version of these measures. Other variables, including owning a computer, graduate versus undergraduate status, previous use of computers, and gender, were also investigated. Statistical analysis of the data revealed there was no significant difference on any variable between the two test administration formats; computer ownership had an effect on both performance and anxiety; anxiety level was higher for graduate students; and females, while starting and ending with lower performance and anxiety levels, made higher gains in performance and decreased their anxiety levels. (Contains 12 references.) (ALF)

Differences Between Computer Administered and Paper Administered Computer Anxiety and Performance Measures

Bob Perkins

School of Education

University of Charleston, SC

Charleston, SC 29424

Phone: (803) 792-5699

INTERNET: PerkinsR@Ashley.CofC.edu

Abstract

This study compares two groups of students using anxiety level and performance as measures. Both groups took a written pretest but one group was administered an anxiety scale (pre and post) and post test on the computer using a HyperCard stack while the other group used paper versions of these measures. Other variables including owning a computer, graduate versus undergraduate, previous use of computers, and gender were also investigated.

Introduction

With the proliferation of computers in education, and especially with computer aided instruction, computer managed instruction and integrated learning systems as instruments of teaching, many factors related to them must be measured. Many schools are investing large amounts of money into Integrated Learning Systems (ILS) as a means of educating students. These systems teach, evaluate, and direct a student's education.

Because hypermedia programs such as HyperCard can be used as authoring systems, classroom teachers are also using these programs more frequently in their classrooms to design lessons. Computerized testing can be an integral part of the Computer-Aided Instruction (CAI) stacks that are created by these teachers. Computers are able to grade, score, and report student grades very quickly, thus freeing the teacher to do more teaching.

If there is a relationship between computer anxiety and performance, and more specifically if the testing is done by computer, the use and design of Integrated Learning Systems as well as classroom hypermedia stacks could be effected. If no relationship is found between anxiety and performance, and if performance is not effected by the evaluating being done by computer, then computer testing should be promoted.

This research is investigating whether computer anxiety is different if the measure is administered by computer than if it is administered by paper and pencil. Also being assessed in this study is a computer administered performance measure versus paper and pencil administration. These comparisons were done to measure the relationship between anxiety and performance. The amount of computer anxiety and performance level exhibited may differ if the measure is administered by computer. Anxiety, if it has an effect on performance, may also be higher if the

performance measure, in the form of a post test, were also administered by computer.

Review of the Literature

The computer band wagon is rolling on faster and faster. With the efforts to quickly reform and improve education and student achievement, school systems are trying to find the optimal way to incorporate the new technologies into the curriculum (Ryan, 1991). Are these new technologies the needed panacea for the ailing educational system? Research on teachers' attitudes related to computers are favorable for incorporation of computers into the curriculum (Dupagne & Krendl, 1992). Computers are seen to have much potential for classroom instruction (Dupagne & Krendl).

These technologies have included stand-alone microcomputers and related peripherals and complex and expensive Integrated Learning Systems (ILS). Many things can effect the success of these investments. Research has shown that factors are: amount of teacher training and amount of administrative commitment (Ryan, 1991). Many Integrated Learning systems incorporate computerized testing into their systems. Immediate feedback, test correction, student diagnosis, standardized presentation, ease of data storage, and records management are all advantages of these systems (Gwinn & Beal, 1987/1988; Davis & Cowles, 1989; Van Horn, 1991).

Many schools do not have the finances to be able to purchase these large computer systems. Many educators use microcomputers offline to create tests (Gwinn & Beal, 1987/1988) with word processing or desktop publishing programs or test generating programs. Microcomputers can also be used to administer tests, grade them, and keep the records current. Many teachers are also incorporating hypermedia-created lessons in their classroom by using them as authoring systems (Johnson, 1991). These hypermedia programs can also be used to create tests that can also be scored by the computer and generate a student progress report.

The variables that effect large Integrated Learning Systems may also effect the students using hypermedia generated tests. Further research on computer testing is necessary because of problems related to previous studies. In a meta-analysis of 40 studies, Ryan (1991) states that it has been difficult to determine variable relationships to student achievement because of under-reporting of study and sample characteristics. Underreported variables in these studies included:

characteristics of color, sound, music, speech generation, and input device (Ryan).

There has been much research involving attitudes toward computers. Both test anxiety and computer anxiety are important concepts when considering computerized testing and they may be related. Most research in the area of attitude relies on self-administered surveys or interviews (Dupagne & Krendl, 1992). Attitudes of prospective teachers, including anxiety, have been found to be related to the computer literacy level of those prospective teachers (Hignite & Echternacht, 1992). Computer anxiety has been a measure in many studies (Reed & Palumbo, 1987/1988; Liu, Reed & Phillips, 1990; Davis & Cowles, 1989; Bandalos & Benson, 1990; Gwinn & Beal, 1987/88; Dimock & Cormier, 1991).

Gwinn & Beal (1987/1988) in their anatomy and physiology classes, used an individualized, self paced instruction and testing system. Most of the students had little or no computer experience prior to these classes. They reported that students in their classes (nursing, allied health, medical assisting, dietetics, and physical education students) using computer administered tests found that general test anxiety using the Mandler and Sarason *Test Anxiety Questionnaire* increased slightly ($p=.001$ for females and .05 for males) from the beginning of the semester to the end. Test anxiety was also found to be negatively correlated with achievement.

Gwinn & Beal (1987/1988) also state that there are two types of test anxiety, facilitating and debilitating test anxiety. They found that paper administered tests initially had higher facilitating and computer administered tests had higher debilitating anxiety. A survey done fourteen weeks later reported that this had inverted, computer administered tests had higher facilitating and paper tests had higher debilitating anxiety.

Computer anxiety can influence test scores (Dimock & Cormier, 1991). This effect could be compounded if the test is administered by computer. Differences have been reported between traditional testing and computerized testing, but Dimock & Cormier state that this could be because of format differences between the tests. Typically, conventional tests present a group of questions at one time while computerized tests only include a single question on the screen at one time. Their research, however, indicated that this was true for an initial test, but a second test with different presentation formats showed no significant differences between the groups.

Computer anxiety may have an effect on student performance when students are required to take tests administered by computer. Studies have shown that computer anxiety is real and measurable, and that it does decrease when the student is involved with learning computer applications or computer literacy (Reed & Palumbo, 1987/1988; Liu, Reed & Phillips, 1990). However, not all students using CAI also use application programs or have taken classes in computer literacy.

Research has also been conducted to determine the effect a computer training module on anxiety (Reed & Palumbo, 1987/1988; Liu, Reed, & Phillips, 1990) using a modified version of Spielberger's *Self-Evaluation Questionnaire* as as a pre and post measure. These studies did not compare administration of the anxiety measure when that measure was administered by computer. If computer anxiety is real and measurable, then it may be more intense if the measurement is conducted by computer.

Computer anxiety has been compared with math anxiety, especially when gender of students enters into the picture (Bandalos & Benson, 1990). It was speculated that these two concepts may be related because of personality variables. Since women have higher levels of math anxiety, their computer anxiety level may also be higher (Bandalos & Benson). It was found that after administering the Computer Anxiety Scale, however, that when they applied a Goodness of Fit Index, that there was no difference between males and females or undergraduate and graduate students enrolled in educational measurements and statistics courses.

Spielberger's State-Trait Personality Inventory was developed to assess three emotions: anxiety, curiosity, and anger (Ben-Zur & Zeidner, 1988). Anxiety is further broken down into two components-state anxiety which is characterized as feelings of tension and is accompanied by increased physiological arousal and trait anxiety which is representative of the frequency and proneness to state anxiety (Ben-Zur & Zeidner). This study sought to compare Israeli students to American students (using Spielberger's data). They found that both cultures reported higher trait anxiety for females (and almost three times as high for Israeli females compared to American females) than males, but that state anxiety showed no difference for gender or culture

Davis & Cowles (1989) reported that students that took tests on computer gave responses that were related to low trait anxiety as measured by Endler's Multidimensional Anxiety Scales and

gave socially desirable answers on the Marlow-Crowne Social Desirability Scale. Students who were administered paper versions of the test did not have this interaction. They speculated that individuals that give socially desirable answers can "fake good" on the scales to invite approval.

Sample

The students in this study (N=83) were from three undergraduate sections and one graduate section of a required class for in-service and pre-service teachers called Computers for Teachers. This is an introductory course that teaches students how to word process, create and use databases, use hypermedia to create instruction, and evaluate software. Students are also exposed to technologies such as CD-ROM, telecommunications, laserdiscs, and video and audio digitizing. Demographic information was collected from the students including gender, major, their computer experience before this class (word processing, database, and computer programming) and whether they own their own computer.

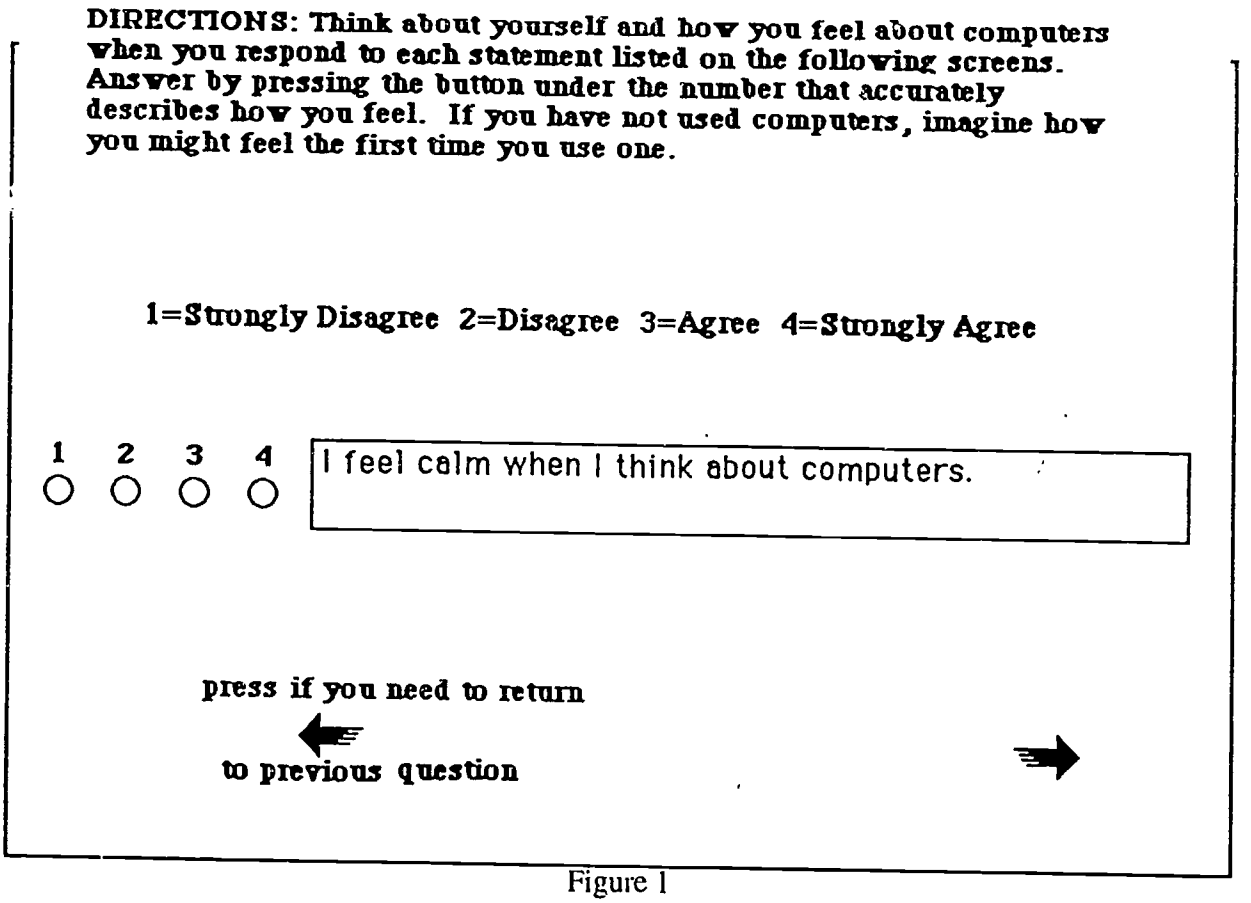

Figure 1

All students were administered a pretest consisting of questions from the previous

semester's exams by paper. One undergraduate and the graduate course also were administered a paper and pencil version of the revised Speilberger's *Self-Evaluation Questionnaire* . The two remaining undergraduate sections were administered the computer version of the revised *Self-Evaluation Questionnaire* on computer using a stack created by the author using *HyperCard* (see Figure 1 for a sample card with the question). The stack asked for the student's name and social security number, and then provided the same written instructions (which stayed on the screen while each question was on the screen) as the paper version of the survey. Students answered the questions by moving the mouse over the selected answer, and clicking the mouse. After the students were finished, the stack closed itself after saving a text file with all of the students' answers and their score. Throughout the stack, the students were allowed to return to previous answers if they wanted to change their answer.

Results

The pre-test was used to determine if the groups were equivalent before the instruction took place in the four classes. An ANOVA revealed no significant difference between the computer and and the paper and pencil group $F(1,81)$, $p=.71$. Based on these results, the computer group (n=44) and non-computer group (n=39) were equivalent at the beginning of the study.

A two factor (group X Anxiety as a repeated measure) ANOVA was performed using pre and post *Self-Evaluation Questionnaire* scores. This anxiety measure can produce scores between 20 and 80 with 80 being the high anxiety level. Since no interaction between group and measure was found $F(1, 81)$, $p= .95$ it was determined that there was no difference between the groups in their level of anxiety either at the pre or post instructional administration of the anxiety measure. There was, however, a significant difference in anxiety apparent, $F(1,81)$, $p=.0001$, as it was reduced from a mean of 48.38 to 37.95.

The the same questions that were administered as a pretest on paper to all students was administered as a post test, but the same group that took the computer anxiety measure took the post test administered by computer. Again, no significant difference between groups was produced, $F(1,81)$, $p=.99$. There was a significant gain from pre to post test $F(1,81)$, $p=.0001$. Students scores went from a mean of 46.7 to 80.04 on a test that had 100 for a maximum score.

Since there were no differences between the computer and non computer groups, the data

was collapsed to compare it to the demographic data. These analyses compared the dependent measures of anxiety and performance against the independent measures of gender, owning a computer, previous computer experience, and undergraduate/graduate status. Other previous studies had mentioned these as potential factors that could effect anxiety and performance.

In comparing those students who owned computers to those who did not on the performance measure, there was a significant difference between the groups, $F(1,81)$, $p=.02$ and also a gain for the repeated measure, $F(1,81)$, $p= 0001$. The students who owned computers ($n=18$) had a higher mean at pre and post test (84.22 and 54.44 respectively) than those that do not own computers ($n=65$) at pre and post test (78.89 and 44.55 respectively). There was no interaction because both groups gained equally on the test.

The same was found to be true for these two groups on the anxiety measure. Those students that owned computers were significantly lower on the anxiety measure, $F(1,81)$, $p=.005$ than those who did not (mean = 41.11 and 50.39 respectively). They also ended the study at a lower anxiety level (mean= 33.11 and 39.11 respectively).

No significant differences were found between undergraduate students ($n=64$) and graduate students ($n=19$) on the performance measure at pre test (mean = 47.70 and 46.38 respectively) and post test (mean = 81.26 and 79.69 respectively). Both groups started equal and ended equal to each other.

Both of these groups also showed no significant differences at the beginning and end of the study on the anxiety measure, but both significantly reduced their anxiety level. The anxiety level of graduate students was slightly higher than undergraduates at the start of the study (mean= 49.53 and 44.32 respectively) and at the end of the study (mean= 38.81 and 35.05, respectively), but not significantly $F(1,81)$, $p=.08$.

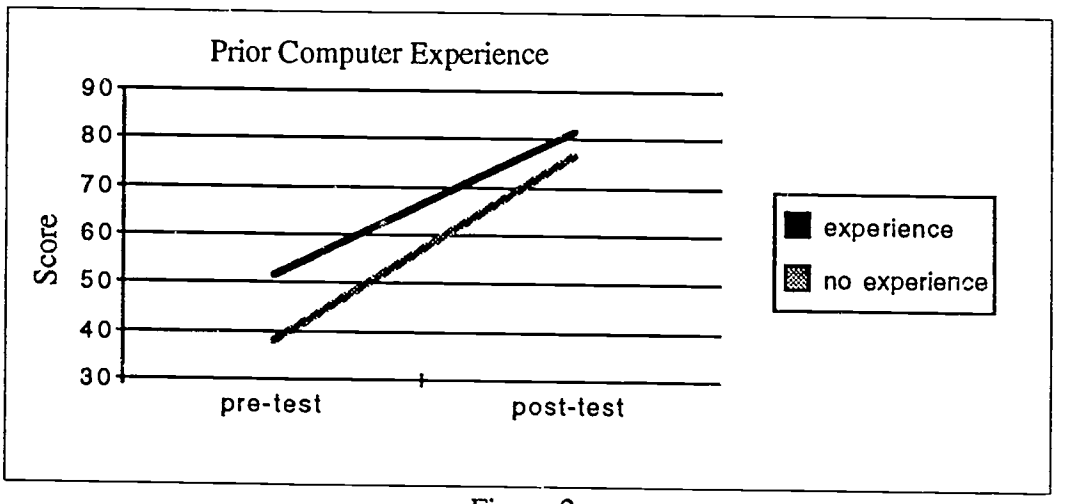

Figure 2

Computer experience for the purpose of this paper was defined as any prior experience using word processing, data base programs, desktop publishing, or computer programming for at least a year. A two factor (prior experience vs. performance) ANOVA using the test as a repeated measure produced an interaction, $F(1, 81)$, $p=.02$ (see Figure 2). Those students that did not have prior experience (n=52) had a lower mean on the pre-test than those who had experience (n=31) (38.26 and 51. 73), a slightly lower post test mean (76.90 and 81.92 respectively), and gained more (38.65 and 30.19 respectively).

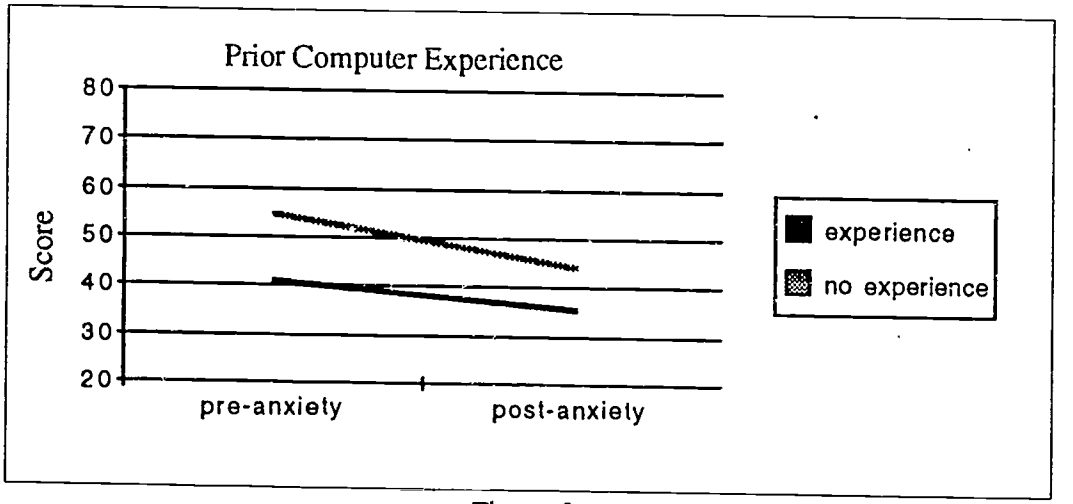

Figure 3

A two factor (prior experience vs. anxiety) ANOVA using the anxiety measure as a repeated measure also produced an interaction, $F(1, 81)$, $p=.02$ (Figure 3). Students without prior

experience initially had higher anxiety mean than those with experience (54.94 and 44.40 respectively), and also finished the study with a higher mean (41.19 and 36.02 respectively). They also made the largest drop in anxiety, however (13.74 and 8.39 respectively).

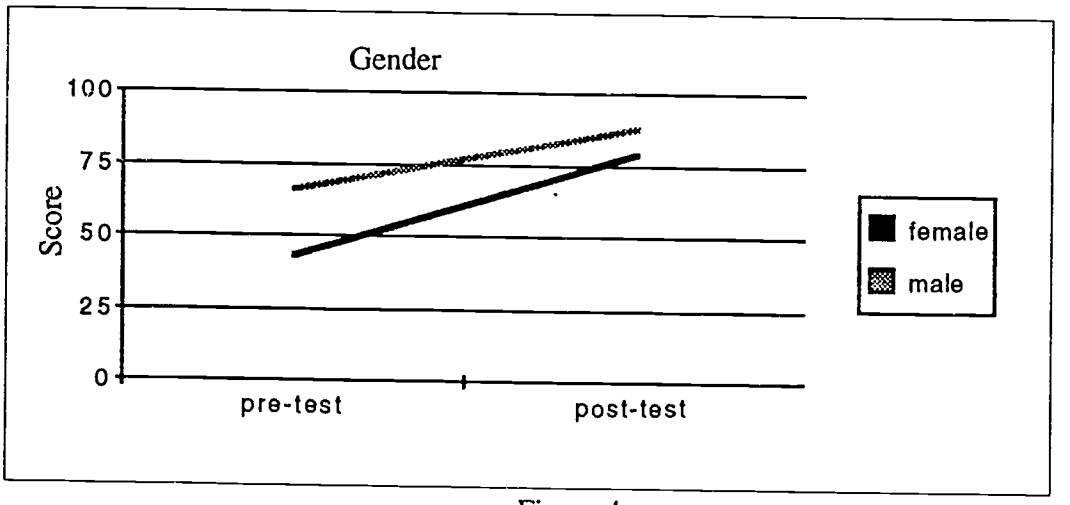

Figure 4

In comparing the test data to gender, a two-factor (gender vs. performance) ANOVA using the test as a repeated measure produced an interaction, $F(1, 81)$, $p=.009$ (Figure 4). There were more females (n=72) than males (n=11) in this study. Females started at a lower mean (43.67 to 66.55) and ended at a lower mean (78.78 to 88.36), but showed a higher gain in mean (35.11 to 21.81).

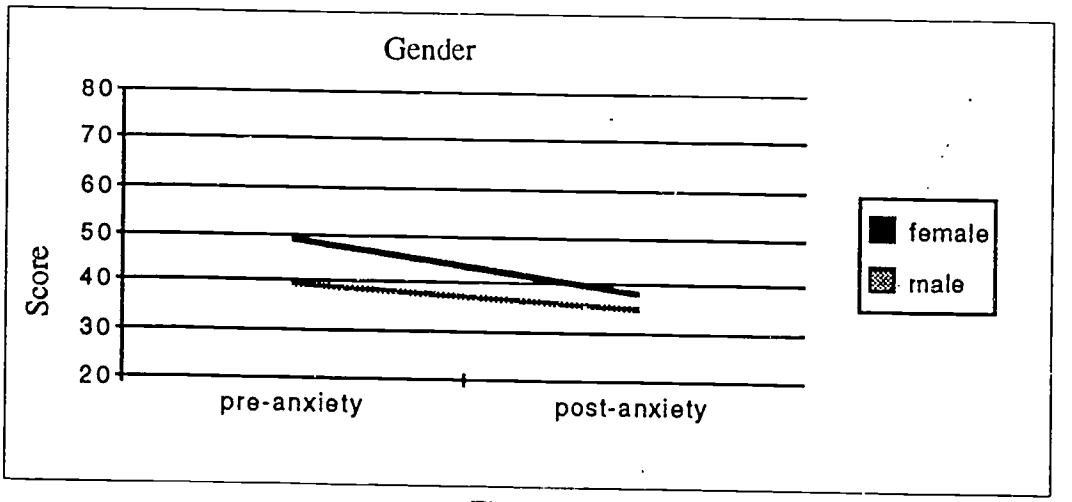

Figure 5

Gender produced significant differences on the anxiety measure, and again, it involved an

interaction (Figure 5). Females exhibited higher anxiety initially in the study (49.71 to 39.64) than males. The means for females were still slightly higher than males at the end of the study (38.38 to 35.18). The anxiety level for females also dropped the most by the end of the study (11.33 to 4.18).

Discussion

The fact that the measures were administered by computer did not significantly differ from the paper and pencil administered test both for computer anxiety and for performance. This would seem to confirm the validity of computerized testing within the classroom using simple hypermedia created tests. This may not generalize to Integrated Learning Systems because the questions were all multiple choice and it was a measure that was not validated. It also measured students' use of application programs as they were using the computer as opposed to CAI a lesson taught by computer.

Another qualification for generalizing the results of the computer administration versus paper and pencil administration is concerned with the students computer experience. Research in the past has shown that student experience with computer application programs reduces the anxiety level (Palumbo & Reed, 1987/1988; Liu, Reed, & Phillips, 1990) of students. Using Integrated Learning Systems is not using true application programs but using Computer-Aided Instruction. The students in this study were involved in using application programs for an entire semester. Students using ILS may only be exposed to CAI for short intervals and may not be getting any "computer experience." This would be worth investigating further.

Those students who owned computers, not surprisingly, did better in performance before and after the study as well as having had a lower anxiety level before and after than those who do not own computers. Apparently, computer ownership, which may translate into computer experience, is a factor to be recognized as having an effect on both performance and anxiety. This would also be a very good area of research in relation to students using ILS programs in their schools.

The results between the undergraduate and graduate students seemed a little surprising since the author taught all sections of these classes. It was not surprising that the groups were initially equal, but during the semester, it was impossible to keep the graduates at the same pace as

the undergraduate students. Graduate students seemed to show more concern for getting good test scores, and so would ask more questions, require more one-on-one instruction, and show less confidence in class. They would then over compensate by studying harder and get higher grades on the same tests during the semester. It always seemed that their anxiety level was higher for graduate students.

There are various possible reasons for these results. The first is that the graduate class had a higher percentage of males than the undergraduate classes. Some of these males seemed to be more knowledgeable about computers initially also. Another factor in relation to the graduate class was that it met once a week for three hours and the undergraduates met either two or three times per week. The graduate students as a whole did not come back to the classroom (which also served as an open computer lab) to do outside of class activities (or practice what they had learned).

Not surprising was the effect that prior computer experience had on both the performance measure and the anxiety measure. Prior studies had established a relationship between both computer anxiety and performance to computer experience, and this just confirmed those results. This however, should suggest that research is necessary to see if this relationship is also true if the performance that is measured is not directly related to the students' experience as this measure was. If computerized testing in a science class, for example, had the same effect (those students with computer experience showing less anxiety and better performance), then this would not be a fair and equitable form of testing.

The one factor that did turn out to be somewhat of a surprise was gender. Females started and ended with lower performance levels, but made higher gains than males, and they also started and ended with higher anxiety levels, but again lowered their anxiety more than males. One thing to consider here is that there were many more females (n=72) than males (n=11). This has been the trend in the computer education classes that the author has taught for the last four years. This would imply that gender may be a factor in testing situations that did not involve the computer in developing experience (literacy). A higher level of anxiety for inexperienced computer using females during computer administered tests could effect the test results.

References

Bandalos, D. & Benson, J. (1990). Testing the factor structure invariance of a computer attitude scale over two grouping conditions. *Educational and Psychological Measurement*, 50(1), 49-60.

Ben-Zur, H. & Zeidner, M. (1988). Sex differences in anxiety, curiosity, and anger: A cross-cultural study. *Sex Roles*, 19(5/6), 335-347.

Davis, C. & Cowles, M. (1989). Automated psychological testing: Method of administration, need for approval, and measures of anxiety. *Educational and Psychological Measurement*, 49(2), 311-320.

Dimock, P. H. & Cormier, P. (1991). The effects of format differences and computer experience on performance and anxiety on a computer-administered test. *Measurement and Evaluation in Counseling and Development*, 24(3), 119-126.

Dupagne, M. & Krendl, K. A. (1992). Teachers' attitudes toward computers: A review of the literature. *Journal on Computing in Education*, 24(3), 420-429.

Gwinn, J. F. & Beal, L. F. (1987/1988). On-line computer testing: Implementation and endorsement. *Journal of Educational Technology Systems*, 16(3), 239-251.

Hignite, M. A. & Echternacht, L. J. (1992). Assessment of the relationships between computer attitudes and computer literacy levels of prospective teachers. *Journal of Research on Computing in Education*, 24(3), 381-391.

Johnson, J. F. (1991). Design assistant: A HyperCard tool for courseware development. *Collegiate Microcomputer*, 9(2), 103-109.

Liu, M., Reed, M. W., & Phillips, P. (1990). Teacher education students and computers: Gender, major, use, occurrence, and anxiety. Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20,

Reed, W. M. & Palumbo, D. (1987/1988). The effects of BASIC instruction on problem solving skills and computer anxiety. Computers in the Schools, 4(3/4), 91-104.

Ryan, A. W. (1991). Meta-analysis of achievement effects of microcomputer applications in elementary schools. *Educational Administration Quarterly*, 27(2), 161-184.

Van Horn, R. (1991). *Advanced technology in education*, Pacific Grove, CA: Brooks/Cole Publishing Company.